US009177387B2

(12) United States Patent
Marks

(10) **Patent No.: US 9,177,387 B2**
(45) **Date of Patent: Nov. 3, 2015**

(54) **METHOD AND APPARATUS FOR REAL TIME MOTION CAPTURE**

(75) Inventor: **Richard L. Marks**, Foster City, CA (US)

(73) Assignee: **Sony Computer Entertainment Inc.**, Tokyo (JP)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 2009 days.

(21) Appl. No.: **10/365,120**

(22) Filed: **Feb. 11, 2003**

(65) **Prior Publication Data**

US 2004/0155962 A1 Aug. 12, 2004

(51) **Int. Cl.**
H04N 5/225 (2006.01)
G06T 7/20 (2006.01)

(52) **U.S. Cl.**
CPC ***G06T 7/2033*** (2013.01); ***G06T 7/2046*** (2013.01); *G06T 2207/10016* (2013.01); *G06T 2207/30196* (2013.01); *G06T 2207/30208* (2013.01)

(58) **Field of Classification Search**
USPC 345/156, 158, 469, 474
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

3,943,277	A	3/1976	Everly et al.	348/19
4,263,504	A	4/1981	Thomas	235/454
4,313,227	A	1/1982	Eder	398/106
4,558,864	A	12/1985	Medwedeff	
4,565,999	A	1/1986	King et al.	345/158
4,787,051	A	11/1988	Olson	364/518
4,802,227	A	1/1989	Elko et al.	381/92
4,823,001	A	4/1989	Kobayashi et al.	356/616
4,843,568	A	6/1989	Krueger et al.	364/518
4,963,858	A	10/1990	Chien	340/710
5,034,986	A	7/1991	Karmann et al.	382/103
5,055,840	A	10/1991	Bartlett	341/31
5,128,671	A	7/1992	Thomas, Jr.	341/20
5,144,594	A	9/1992	Gilchrist	367/129
5,195,179	A	3/1993	Tokunaga	395/161
5,260,556	A	11/1993	Lake et al.	235/494

(Continued)

FOREIGN PATENT DOCUMENTS

EP	0 353 200 A2	1/1990		
EP	0353200	1/1990	....	G01S 17/10

(Continued)

OTHER PUBLICATIONS

"The Tracking Cube: A Three-Dimensional Input Device," IBM Technical Disclosure Bulletin, Aug. 1, 1989, pp. 91-95, vol. 32, No. 3B, IBM Corp. New York, US.

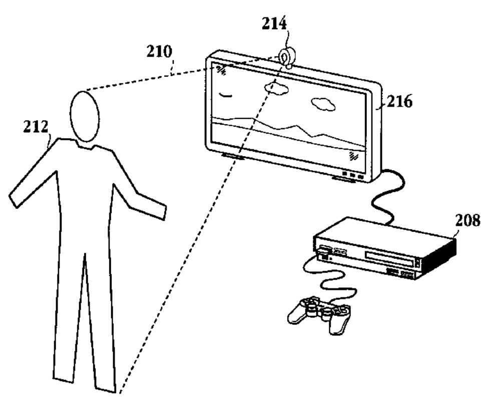
(Continued)

Primary Examiner — William Boddie
Assistant Examiner — Bryan Earles
(74) *Attorney, Agent, or Firm* — Martine Penilla Group, LLP

(57) **ABSTRACT**

Methods for real time motion capture for control of a video game character is provided. In one embodiment, a method initiates with defining a model of a control object. Then, a location of a marker on the model is identified. Next, movement associated with the control object is captured. Then, the movement associated with the control object is interpreted to change a position of the model. Next, movement of the character being presented on the display screen is controlled according to the change of position of the model. A computer readable media and a processing system enabling control of video character through real time motion capture are also provided.

14 Claims, 11 Drawing Sheets



(56) **References Cited**

U.S. PATENT DOCUMENTS

5,297,061	A *	3/1994	Dementhon et al.	345/180
5,335,011	A	8/1994	Addeo et al.	348/15
5,394,168	A	2/1995	Smith, III et al.	345/156
5,426,450	A	6/1995	Drumm	345/168
5,435,554	A	7/1995	Lipson	273/88
5,453,758	A	9/1995	Sato	345/158
5,455,685	A	10/1995	Mori	348/363
5,473,701	A	12/1995	Cezanne et al.	381/92
5,485,273	A	1/1996	Mark et al.	356/350
5,517,333	A	5/1996	Tamura et al.	358/158
5,528,265	A	6/1996	Harrison	345/158
5,534,917	A	7/1996	MacDougall	348/169
5,543,818	A	8/1996	Scott	345/168
5,557,684	A	9/1996	Wang et al.	382/107
5,563,988	A	10/1996	Maes et al.	345/421
5,568,928	A	10/1996	Munson et al.	463/37
5,581,276	A	12/1996	Cipolla et al.	345/156
5,583,478	A	12/1996	Renzi	340/407.1
5,586,231	A	12/1996	Florent et al.	
5,611,000	A	3/1997	Szeliski et al.	
5,611,731	A	3/1997	Bouton et al.	463/37
5,616,078	A *	4/1997	Oh	463/8
5,638,228	A	6/1997	Thomas, III	360/60
5,649,021	A	7/1997	Matey et al.	382/128
5,675,825	A	10/1997	Dreyer et al.	395/800
5,677,710	A	10/1997	Thompson-Rohrlich	345/173
5,706,364	A	1/1998	Kopec et al.	382/159
5,768,415	A	6/1998	Jagadish et al.	382/154
5,796,354	A	8/1998	Cartabiano et al.	341/22
5,818,424	A	10/1998	Korth	345/158
5,846,086	A	12/1998	Bizzi et al.	434/247
5,850,222	A	12/1998	Cone	345/418
5,850,473	A	12/1998	Andersson	
5,861,910	A	1/1999	McGarry et al.	348/87
5,870,100	A	2/1999	DeFreitas	345/441
5,883,616	A	3/1999	Koizumi et al.	345/156
5,889,672	A	3/1999	Schuler et al.	700/83
5,900,863	A	5/1999	Numazaki	345/158
5,913,727	A	6/1999	Ahdoot	463/39
5,914,723	A	6/1999	Gajewska	345/597
5,917,493	A	6/1999	Tan et al.	715/835
5,917,936	A	6/1999	Katto	382/154
5,923,318	A	7/1999	Zhai et al.	
5,929,444	A	7/1999	Leichner	250/341.7
5,930,383	A	7/1999	Netzer	382/154
5,930,741	A	7/1999	Kramer	702/153
5,937,081	A	8/1999	O'Brill et al.	382/111
5,959,596	A	9/1999	McCarten et al.	345/2
5,963,250	A	10/1999	Parker et al.	348/211.6
5,978,722	A	11/1999	Takasuka et al.	705/16
5,993,314	A	11/1999	Dannenberg et al.	461/1
6,009,210	A	12/1999	Kang	382/276
6,014,167	A	1/2000	Suito et al.	348/169
6,021,219	A	2/2000	Andersson et al.	382/190
6,031,934	A	2/2000	Ahmad et al.	382/154
6,037,942	A	3/2000	Millington	715/835
6,044,181	A	3/2000	Szeliski et al.	382/284
6,049,619	A	4/2000	Anandan et al.	382/107
6,056,640	A *	5/2000	Schaaij	463/4
6,057,909	A	5/2000	Yahav et al.	356/5.04
6,061,055	A	5/2000	Marks	382/276
6,072,494	A	6/2000	Nguyen	715/863
6,075,895	A *	6/2000	Qiao et al.	382/218
6,091,905	A	7/2000	Yahav et al.	396/106
6,097,369	A	8/2000	Wambach	
6,100,517	A	8/2000	Yahav et al.	250/208.1
6,101,289	A	8/2000	Kellner	382/276
6,115,052	A *	9/2000	Freeman et al.	345/473
6,134,346	A	10/2000	Berman et al.	382/163
6,144,367	A	11/2000	Berstis	345/158
6,151,009	A	11/2000	Kanade et al.	345/113
6,157,368	A	12/2000	Fager	345/156
6,160,540	A	12/2000	Fishkin et al.	345/184
6,173,059	B1	1/2001	Huang et al.	381/92
6,184,863	B1	2/2001	Sibert et al.	345/156
6,191,773	B1	2/2001	Maruno et al.	345/158
6,195,104	B1	2/2001	Lyons	345/473
6,215,898	B1	4/2001	Woodfill et al.	382/154
6,243,074	B1	6/2001	Fishkin et al.	
6,243,491	B1	6/2001	Andersson	382/165
6,275,213	B1	8/2001	Tremblay et al.	345/156
6,281,930	B1	8/2001	Parker et al.	348/211.9
6,297,838	B1	10/2001	Chang et al.	715/863
6,307,549	B1	10/2001	King et al.	715/810
6,307,568	B1	10/2001	Rom	345/629
6,323,839	B1	11/2001	Fukuda et al.	345/157
6,323,942	B1	11/2001	Bamji	356/5.01
6,326,901	B1	12/2001	Gonzales	340/7.2
6,327,073	B1	12/2001	Yahav et al.	359/321
6,331,911	B1	12/2001	Manassen et al.	359/260
6,346,929	B1	2/2002	Fukushima et al.	354/8
6,351,661	B1	2/2002	Cosman	600/426
6,371,849	B1	4/2002	Togami	463/4
6,375,572	B1	4/2002	Masuyama et al.	463/43
6,392,644	B1	5/2002	Miyata et al.	345/419
6,393,142	B1	5/2002	Swain et al.	382/154
6,394,897	B1	5/2002	Togami	463/4
6,400,374	B2	6/2002	Lanier	345/630
6,411,392	B1	6/2002	Bender et al.	358/1.14
6,411,744	B1	6/2002	Edwards	382/294
6,417,836	B1	7/2002	Kumar et al.	345/156
6,441,825	B1	8/2002	Peters	345/473
6,473,516	B1	10/2002	Kawaguchi et al.	382/100
6,498,860	B1	12/2002	Sasaki et al.	382/103
6,513,160	B2	1/2003	Dureau	725/9
6,519,359	B1	2/2003	Nafis et al.	382/154
6,533,420	B1	3/2003	Eichenlaub	353/7
6,542,927	B2	4/2003	Rhoads	709/217
6,545,706	B1	4/2003	Edwards et al.	348/169
6,556,704	B1	4/2003	Chen	382/154
6,577,748	B2	6/2003	Chang	382/100
6,580,414	B1	6/2003	Wergen et al.	345/156
6,580,415	B1	6/2003	Kato et al.	345/156
6,587,573	B1	7/2003	Stam et al.	382/104
6,587,835	B1	7/2003	Treyz et al.	705/14
6,593,956	B1	7/2003	Potts et al.	348/14.09
6,595,642	B2	7/2003	Wirth	351/211
6,621,938	B1	9/2003	Tanaka et al.	382/276
6,661,914	B2	12/2003	Dufour	382/154
6,674,415	B2	1/2004	Nakamura et al.	345/32
6,677,967	B2	1/2004	Sawano et al.	715/839
6,677,987	B1	1/2004	Girod	348/171
6,709,108	B2	3/2004	Levine et al.	351/211
6,712,703	B2 *	3/2004	Miyamoto et al.	463/43
6,720,949	B1	4/2004	Pryor et al.	
6,727,988	B2	4/2004	Kim et al.	356/319
6,741,741	B2	5/2004	Farrell	382/199
6,746,124	B2	6/2004	Fischer et al.	353/43
6,751,338	B1	6/2004	Wallack	382/106
6,753,849	B1	6/2004	Curran et al.	345/158
6,769,769	B2	8/2004	Podlleanu et al.	351/221
6,772,057	B2	8/2004	Breed et al.	701/45
6,774,939	B1	8/2004	Peng	348/231.4
6,785,329	B1	8/2004	Pan et al.	375/240.08
6,789,967	B1	9/2004	Forester	400/489
6,791,531	B1	9/2004	Johnston et al.	345/157
6,795,068	B1	9/2004	Marks	
6,809,776	B1	10/2004	Simpson et al.	348/565
6,819,318	B1	11/2004	Geng	345/420
6,847,311	B2	1/2005	Li	341/28
6,870,526	B2	3/2005	Zngf et al.	345/185
6,873,747	B2	3/2005	Askary	382/295
6,881,147	B2	4/2005	Naghi et al.	463/35
6,917,688	B2	7/2005	Yu et al.	381/94.7
6,919,824	B2	7/2005	Lee	341/20
6,924,787	B2	8/2005	Kramer et al.	345/156
6,928,180	B2	8/2005	Stam et al.	382/104
6,931,125	B2	8/2005	Smallwood	379/433.07
6,943,776	B2	9/2005	Ehrenburg	345/168
6,947,576	B2	9/2005	Stam et al.	382/104
6,951,515	B2	10/2005	Ohshima et al.	463/31
6,952,198	B2	10/2005	Hansen	345/158
6,970,183	B1	11/2005	Monroe	348/143
6,990,639	B2	1/2006	Wilson	715/863

(56) **References Cited**

U.S. PATENT DOCUMENTS

7,016,411	B2	3/2006	Azuma et al.	375/240.08
7,016,532	B2	3/2006	Boncyk et al.	382/165
7,039,199	B2	5/2006	Rui	381/92
7,039,253	B2	5/2006	Matsuoka et al.	382/295
7,042,440	B2	5/2006	Pryor et al.	345/158
7,043,056	B2	5/2006	Edwards et al.	382/103
7,054,452	B2	5/2006	Ukita	381/92
7,059,962	B2	6/2006	Watashiba	463/2
7,061,507	B1	6/2006	Tuomi et al.	345/611
7,098,891	B1	8/2006	Pryor	345/158
7,102,615	B2	9/2006	Marks	345/156
7,106,366	B2	9/2006	Parker et al.	348/222.1
7,113,635	B2	9/2006	Robert et al.	382/154
7,116,330	B2	10/2006	Marshall et al.	345/474
7,139,767	B1 *	11/2006	Taylor et al.	707/102
7,148,922	B2	12/2006	Shimada	348/370
7,156,311	B2	1/2007	Attia et al.	235/462.46
7,158,118	B2	1/2007	Liberty	345/158
7,161,634	B2	1/2007	Long	384/624
7,164,413	B2	1/2007	Davis et al.	345/163
7,174,312	B2	2/2007	Harper et al.	705/35
7,183,929	B1	2/2007	Antebi et al.	340/573.1
7,212,308	B2	5/2007	Morgan	358/1.18
7,223,173	B2	5/2007	Masuyama et al.	463/36
7,224,384	B1	5/2007	Iddan et al.	348/207.99
7,227,526	B2	6/2007	Hildreth et al.	345/156
7,227,976	B1	6/2007	Jung et al.	382/103
7,239,301	B2	7/2007	Liberty et al.	345/158
7,245,273	B2	7/2007	Eberl et al.	351/211
7,259,375	B2	8/2007	Tichit et al.	250/341.8
7,262,760	B2	8/2007	Liberty	345/158
7,263,462	B2	8/2007	Funge et al.	702/179
7,274,305	B1	9/2007	Luttrell	340/870.02
7,277,526	B2	10/2007	Rifkin et al.	345/156
7,283,679	B2	10/2007	Okada et al.	382/260
7,296,007	B1	11/2007	Funge et al.	706/47
7,301,530	B2	11/2007	Lee et al.	345/158
7,301,547	B2	11/2007	Martins et al.	345/633
7,305,114	B2	12/2007	Wolff et al.	709/200
7,346,387	B1	3/2008	Wachter et al.	600/476
7,352,359	B2	4/2008	Zalewski et al.	345/156
7,364,297	B2	4/2008	Goldfain et al.	351/206
7,379,559	B2	5/2008	Wallace et al.	382/100
7,391,409	B2	6/2008	Zalewski et al.	345/156
7,436,887	B2	10/2008	Yeredor et al.	375/240
7,446,650	B2	11/2008	Schofield et al.	340/425.5
7,489,299	B2	2/2009	Liberty et al.	345/163
7,545,926	B2	6/2009	Mao	379/406.08
7,555,157	B2	6/2009	Davidson et al.	382/154
7,558,698	B2	7/2009	Funge et al.	702/179
7,613,610	B1	11/2009	Zimmerman et al.	704/235
7,623,115	B2	11/2009	Marks	345/156
7,627,139	B2	12/2009	Marks et al.	382/103
7,636,645	B1	12/2009	Yen et al.	702/152
7,636,697	B1	12/2009	Dobson et al.	706/12
7,636,701	B2	12/2009	Funge et al.	706/47
7,646,372	B2	1/2010	Marks et al.	345/156
7,665,041	B2	2/2010	Wilson et al.	715/860
7,697,700	B2	4/2010	Mao	381/94.3
7,721,231	B2	5/2010	Wilson	715/863
2001/0056477	A1 *	12/2001	McTernan et al.	709/219
2002/0021277	A1	2/2002	Kramer et al.	
2002/0023027	A1	2/2002	Simonds	705/26
2002/0036617	A1	3/2002	Pryor	345/156
2002/0085097	A1	7/2002	Colmenarez et al.	348/211
2002/0134151	A1	9/2002	Naruoka et al.	73/291
2002/0158873	A1	10/2002	Williamson	345/427
2003/0020718	A1	1/2003	Marshall et al.	
2003/0022716	A1	1/2003	Park et al.	463/36
2003/0032466	A1	2/2003	Watashiba	
2003/0093591	A1	5/2003	Hohl	710/22
2003/0100363	A1	5/2003	Ali	436/30
2003/0107500	A1	6/2003	Lee	
2003/0123705	A1	7/2003	Stam et al.	382/104
2003/0160862	A1	8/2003	Charlier et al.	348/14.08
2003/0184451	A1	10/2003	Li	
2004/0001082	A1	1/2004	Said	345/730
2004/0017355	A1	1/2004	Shim	345/157
2004/0035925	A1	2/2004	Wu et al.	235/380
2004/0063480	A1 *	4/2004	Wang	463/8
2004/0063481	A1	4/2004	Wang	463/8
2004/0070565	A1	4/2004	Nayar et al.	345/156
2004/0087366	A1	5/2004	Shum et al.	463/36
2004/0095327	A1	5/2004	Lo	345/169
2004/0140955	A1	7/2004	Metz	345/166
2004/0178576	A1	9/2004	Hillis et al.	273/148
2004/0189720	A1	9/2004	Wilson et al.	345/863
2004/0212589	A1	10/2004	Hall et al.	345/156
2004/0213419	A1	10/2004	Varma et al.	381/92
2004/0227725	A1	11/2004	Calarco et al.	345/156
2004/0254017	A1	12/2004	Cheng et al.	463/35
2005/0037844	A1	2/2005	Shum et al.	463/36
2005/0047611	A1	3/2005	Mao	381/94.7
2005/0088369	A1	4/2005	Yoshioka	345/60
2005/0105777	A1	5/2005	Kozlowski et al.	382/115
2005/0117045	A1	6/2005	Abdellatif et al.	384/335
2005/0162385	A1	7/2005	Doi et al.	345/156
2005/0226431	A1	10/2005	Mao	381/61
2006/0033713	A1	2/2006	Pryor	345/158
2006/0035710	A1	2/2006	Festejo et al.	463/36
2006/0038819	A1	2/2006	Festejo et al.	345/530
2006/0204012	A1	9/2006	Marks et al.	381/26
2006/0233389	A1	10/2006	Mao et al.	381/92
2006/0252541	A1	11/2006	Zalewski et al.	463/156
2006/0252543	A1	11/2006	Van Noland et al.	463/37
2006/0256081	A1	11/2006	Zalewski et al.	345/156
2006/0264258	A1	11/2006	Zalewski et al.	463/36
2006/0264259	A1	11/2006	Zalewski et al.	463/36
2006/0264260	A1	11/2006	Zalewski et al.	463/36
2006/0269072	A1	11/2006	Mao	381/56
2006/0269073	A1	11/2006	Mao	381/56
2006/0274032	A1	12/2006	Mao et al.	345/156
2006/0274911	A1	12/2006	Mao et al.	381/334
2006/0280312	A1	12/2006	Mao	381/56
2006/0282873	A1	12/2006	Zalewski et al.	725/133
2006/0287084	A1	12/2006	Mao et al.	463/37
2006/0287085	A1	12/2006	Mao et al.	463/37
2006/0287086	A1	12/2006	Zalewski et al.	436/37
2006/0287087	A1	12/2006	Zalewski et al.	463/37
2007/0015559	A1	1/2007	Zalewski et al.	463/1
2007/0021208	A1	1/2007	Mao et al.	463/36
2007/0025562	A1	2/2007	Zalewski et al.	381/92
2007/0061413	A1	3/2007	Larsen et al.	709/217
2007/0066394	A1	3/2007	Ikeda et al.	463/37
2007/0072675	A1	3/2007	Hammano et al.	463/42
2007/0120834	A1	5/2007	Boillot	354/103
2007/0120996	A1	5/2007	Boillot	384/345
2007/0260340	A1	11/2007	Mao	700/94
2007/0260517	A1	11/2007	Zalewski et al.	705/14
2007/0261077	A1	11/2007	Zalewski et al.	725/35
2008/0056561	A1	3/2008	Sawachi	382/154
2008/0070684	A1	3/2008	Haigh-Hutchinson	463/32
2008/0091421	A1	4/2008	Gustavsson	704/233
2008/0208613	A1	8/2008	Scibora	705/1
2009/0010494	A1	1/2009	Bechtel et al.	382/104
2009/0016642	A1	1/2009	Hart	382/278
2009/0221368	A1	9/2009	Yen et al.	463/32
2009/0221374	A1	9/2009	Yen et al.	463/42
2009/0288064	A1	11/2009	Yen et al.	717/106
2010/0004896	A1	1/2010	Yen et al.	702/153
2010/0137064	A1	6/2010	Shum et al.	463/36

FOREIGN PATENT DOCUMENTS

EP	0 613 294	A1	8/1994	
EP	0652686		5/1995	H04R 1/40
EP	0 750 202	A1	12/1996	
EP	0750202		12/1996	G01S 17/00
EP	0 823 683	A1	2/1998	
EP	0835676		4/1998	A63F 9/22
EP	1132888	A1 *	9/2001	
EP	1098686		5/2003	A63F 13/04
EP	1435258		7/2004	A63F 13/04
FR	2814965		4/2002	A63H 30/00

(56) **References Cited**

FOREIGN PATENT DOCUMENTS

FR	2 832 892	5/2003	
GB	2206716	11/1989	G06K 7/12
GB	2376397	11/2002	G06F 3/00
GB	2388418	11/2003	G06F 3/033
JP	1284897	11/1989	G10H 1/34
JP	6102980	4/1994	G06F 3/02
JP	9128141	5/1997	G06F 3/033
JP	9185456	7/1997	G06F 3/033
JP	1138949	2/1999	G09G 5/00
JP	2000-172431	6/2000	G06F 3/033
JP	2001-166676	6/2001	G09B 9/00
JP	2004-145448	5/2004	G06T 17/40
WO	WO 88/05942	8/1988	G06F 3/033
WO	WO 98/48571	10/1998	H04N 5/45
WO	WO 99/26198	10/1999	
WO	WO 00/23870	4/2000	
WO	WO 01/18563 A1	3/2001	
WO	WO 02/27456	2/2002	G06F 3/00
WO	WO 3/079179	9/2003	G06F 3/033
WO	WO 2005/073838	8/2005	G06F 3/033
WO	WO 2008/056180	5/2008	G01B 11/04

OTHER PUBLICATIONS

Jojie et al., Tracking Self-Occluding Articulated Objects in Dense Disparity Maps, Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on Kerkyra, Greece Sep. 20-27, 1999, Los Alamitos, CA, USA, IEEE Comput. Soc, US, Sep. 20, 1999,pp. 123-130.

Iddan et al. "*3D Imaging in the Studio* (*and Elsewhere . . .*)", Proceedings of the SPIE, SPIE, Bellingham, VA, US, vol. 4298, Jan. 24, 2001, pp. 48-55, XP008005351.

K. B. Shimoga, et al., "Touch and Force Reflection for Telepresence Surgery", Engineering in Medicine and Biology Opportunities of the IEEEE, Baltimore, MD, USA, Nov. 3, 1994, New York, New York, USA, pp. 1049-1050.

Klinker, et al., "Distributed User Tracking Concepts for Augmented Reality Applications", pp. 37-44, Augmented Reality, 2000, IEEE and ACM Int'l Symposium, Oct. 2000, XP010520308, ISBN: 0-7695-0846-4, Germany.

Nakagawa, et al., "A Collision Detection and Motion Image Synthesis Between a Background Image and a Foreground 3-Dimensional Object", TVRSJ Bol. 4, No. 2, pp. 425-430, 1999, Japan.

Mihara, et al., "A Realtime Vision-Based Interface Using Motion Processor and Applications to Robotics", vol. J84-D-11, No. 9, pp. 2070-2078, Sep. 2001, Japan.

Nakamura, et al., "A Consideration on Reconstructing 3-D Model Using Object Views", 2004-01601-003, pp. 17-21, Kokkaido University, Japan, nakamura@media.eng.hokudai.ac.jp.

Nishida, et al., "A Method of Estimating Human Shapes by Fitting the Standard Human Model to Partial Measured Data", D-II vol. J84-D-II, No. 7, pp. 1310-1318, Jul. 2001.

Wilson & Darrell, "Audio-Video Array Source Localization for Intelligent Environments", 2002 IEEE Dept. of Electrical Eng and Computer Science, MIT, Cambridge, MA 02139.

Fiala, et al., "A Panoramic Video and Acoustic Beamforming Sensor for Videoconferencing", 2004 IEEE, Computational Video Group, National Research Council, Ottawa, Canada K1A 0R6.

Hemmi, et al., "3-D Natural Interactive Interface-Using Marker Tracking from a Single View",Sep. 9, 1991, Systems and Computers in Japan.

Kanade, et al., "A Stereo Machine for Video-rate Dense Depth Mapping and Its New Application" 1996. CVPR 96, IEEE Computer Society Conference, pp. 196-202 (022).

Gvili, et al., "Depth Keying", SPIE vol. 5006 (2003), 2003 SPIE-IS &T, pp. 564-574 (031).

Bolt, R.A., "Put-that-there": voice and gesture at the graphics interface,Computer Graphics, vol. 14, No. 3 (ACM SIGGRAPH Conference Proceedings) Jul. 1980, pp. 262-270.

DeWitt, Thomas and Edelstein, Phil "Pantomation: A System for Position Tracking", Proceedings of the 2^{nd} Symposium on Small Computers in the Arts, Oct. 1982, pp. 61-69.

Ephraim et al. "Speech Enhancement Using a Minimum Mean-Square Error Log-Spectral Amplitude Estimator", 1985, IEEE.

Ephraim et al. "Speech Enhancement Using a Minimum Mean-Square Error Short-Time Spectral Amplitude Estimator", 1984, IEEE.

Richardson et al. "Virtual Network Computing", 1998, IEEE Internet Computing vol. 2.

XP-002453974, "CFS and FS95/98/2000: How to Use the Trim Controls to Keep Your Aircraft Level", Aug. 10, 2007, http://support.microsof/com/?scid=kb%3Ben-us%3B175195&x=13&y=15.

* cited by examiner

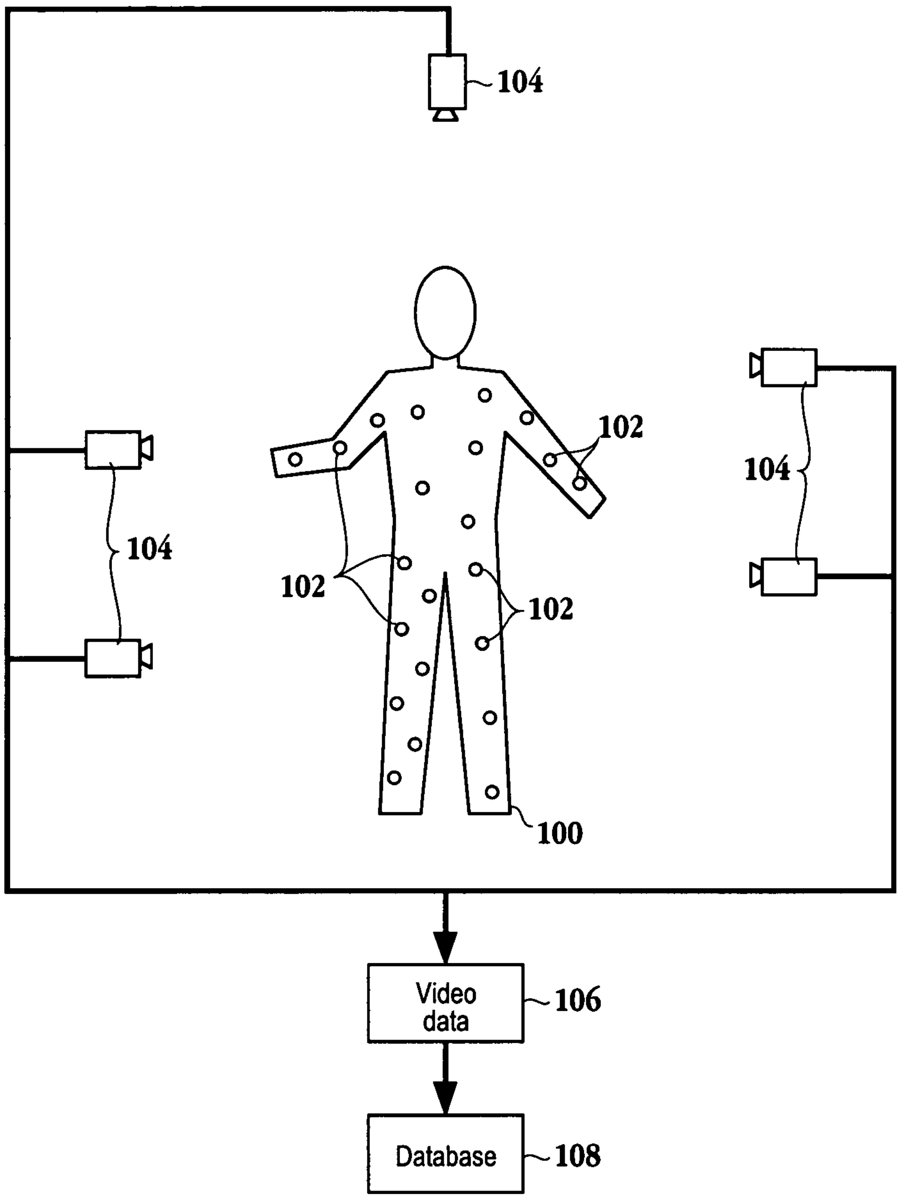


Fig. 1 *(prior art)*

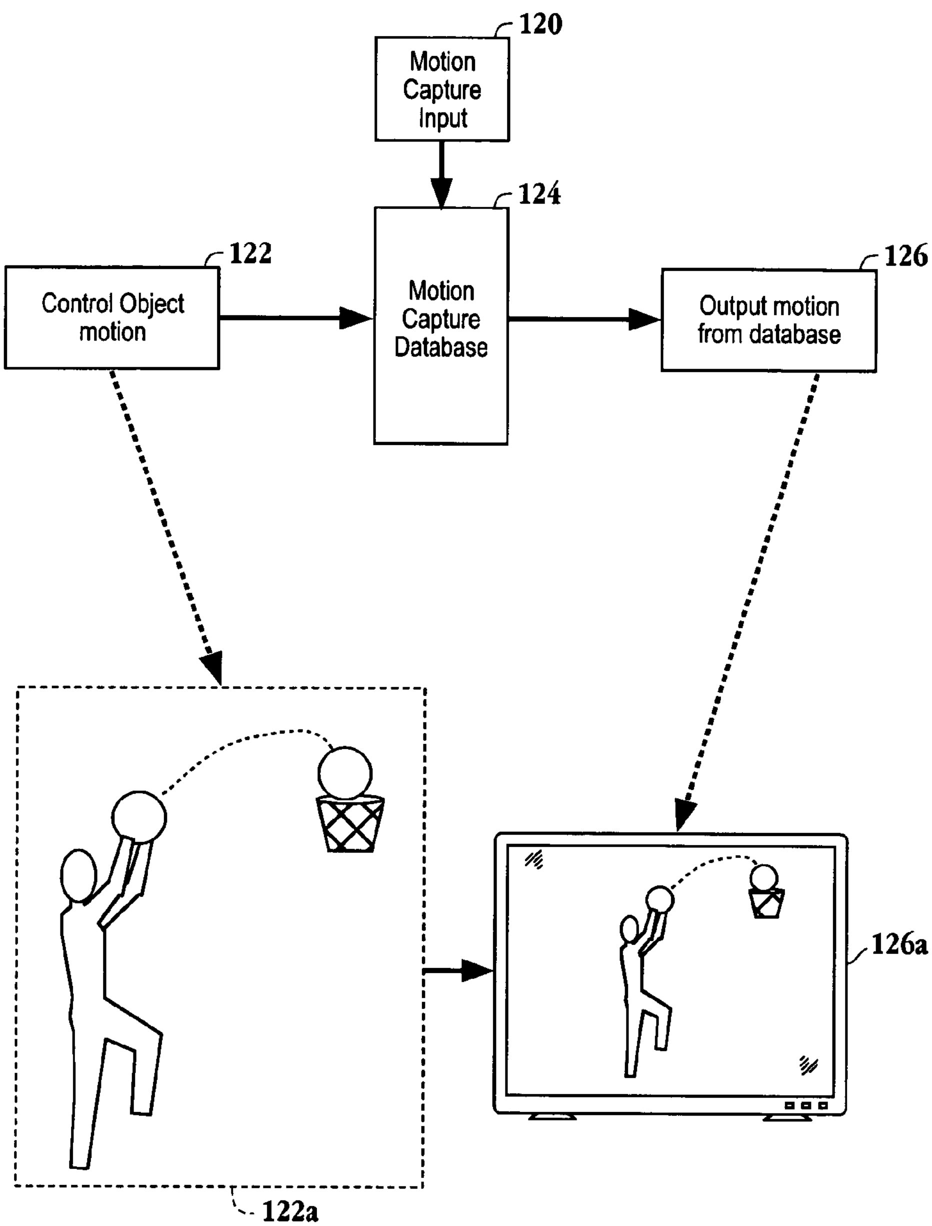

120
Motion Capture Input
124
Motion Capture Database
122
Control Object motion
126
Output motion from database
126a
122a

Fig. 2

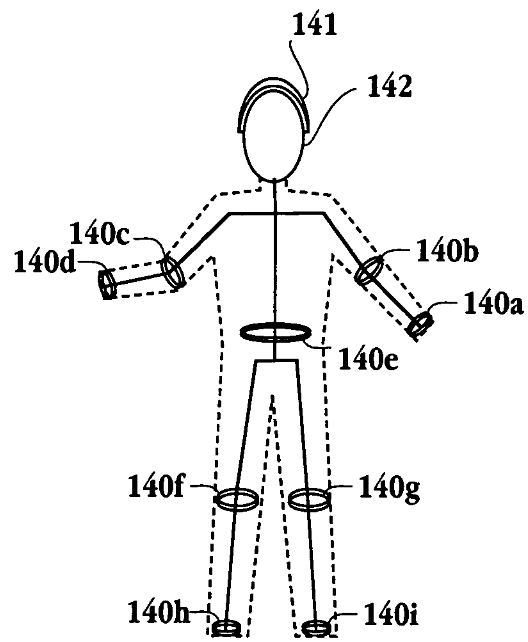


Fig. 3

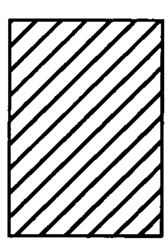


Fig. 4A

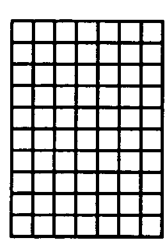


Fig. 4B

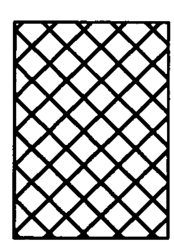


Fig. 4C

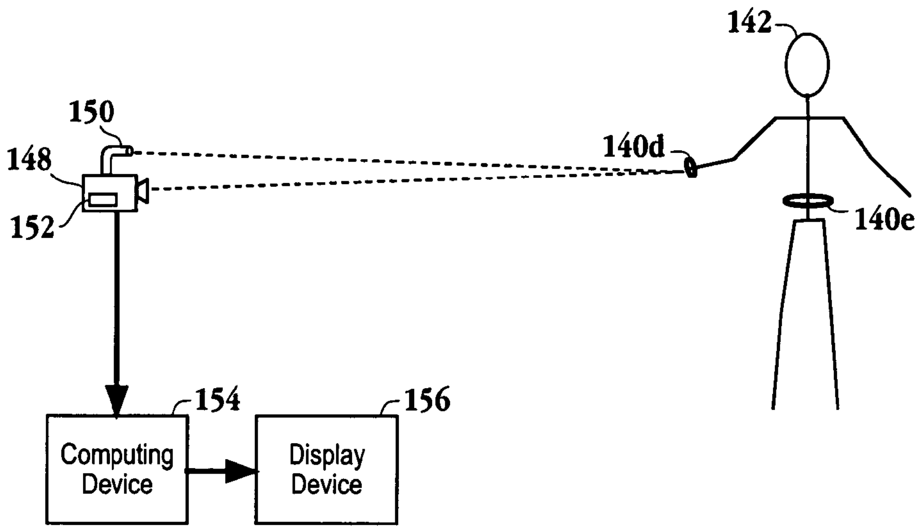

142
150
148
140d
152
140e
154
156
Computing Device
Display Device

Fig. 5

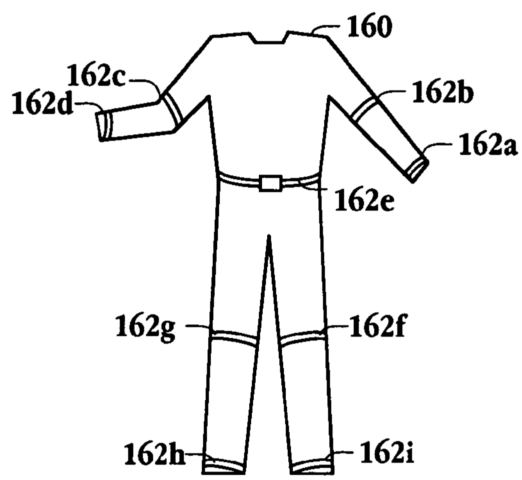

160
162c
162d
162b
162a
162e
162g
162f
162h
162i

Fig. 6A

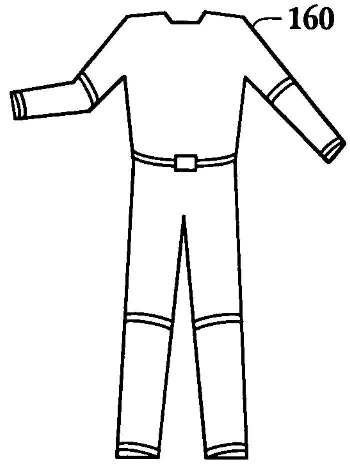


Fig. 6B

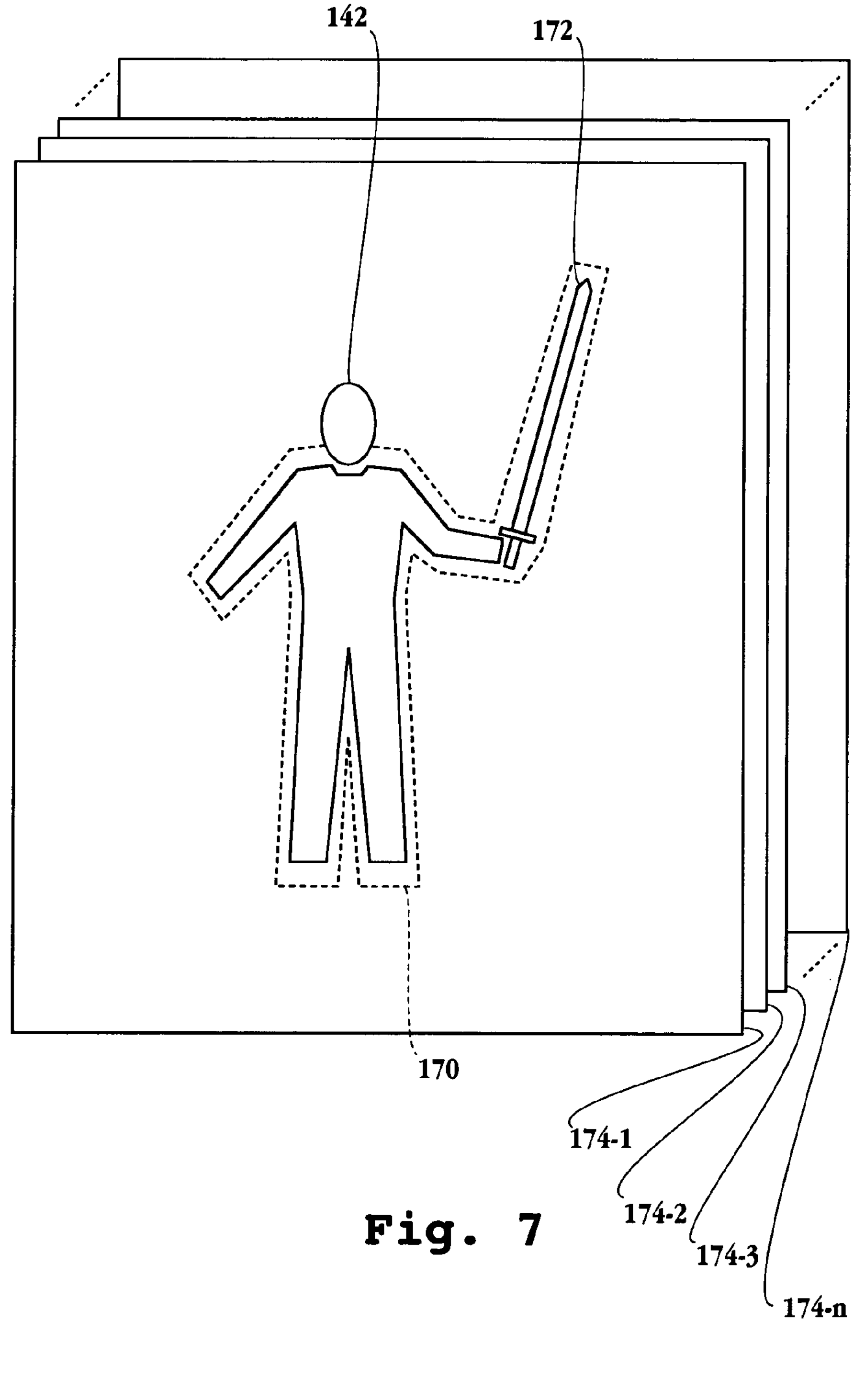


Fig. 7

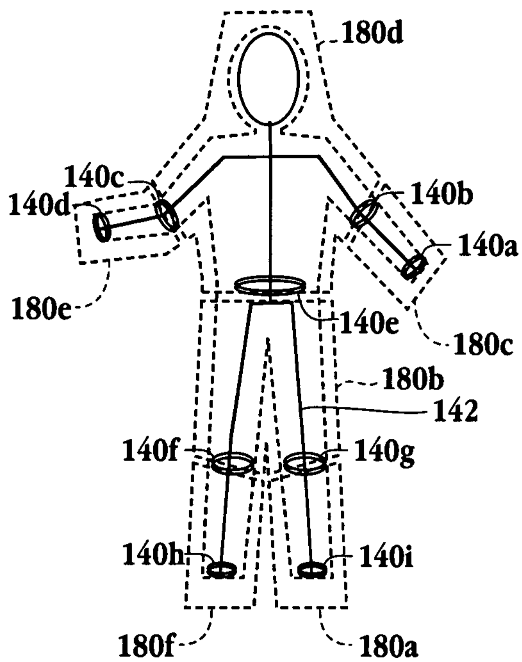

180d
140c
140b
140d
140a
180e
140e
180c
180b
142
140f
140g
140h
140i
180f
180a

Fig. 8

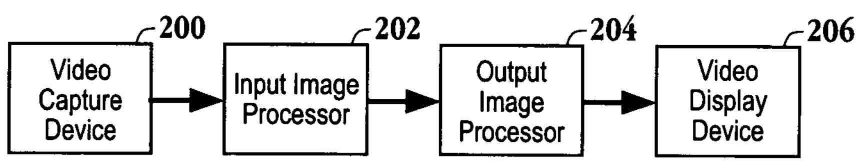

200
Video Capture Device
202
Input Image Processor
204
Output Image Processor
206
Video Display Device

Fig. 9

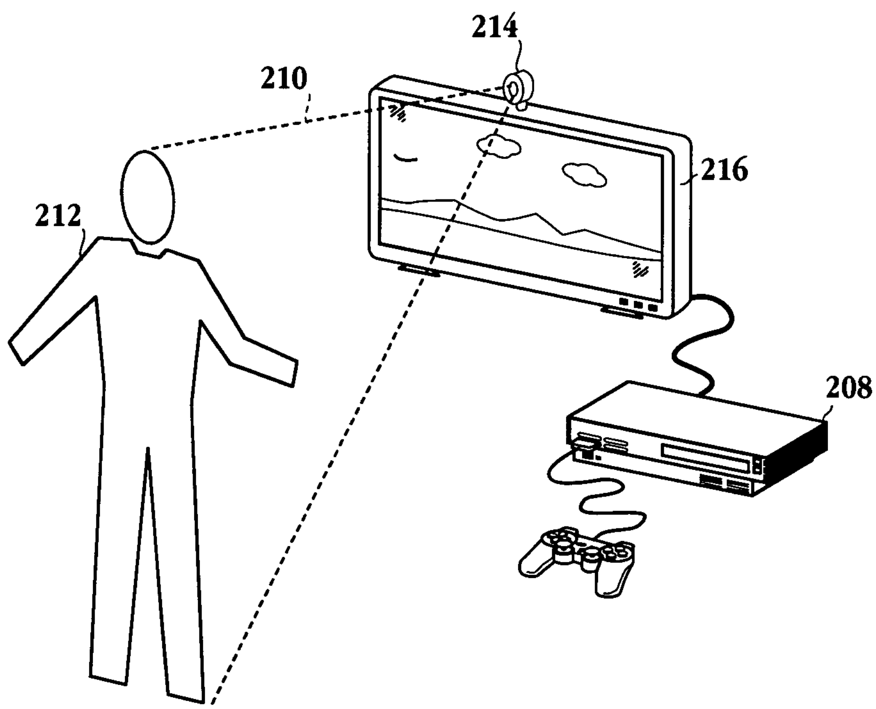
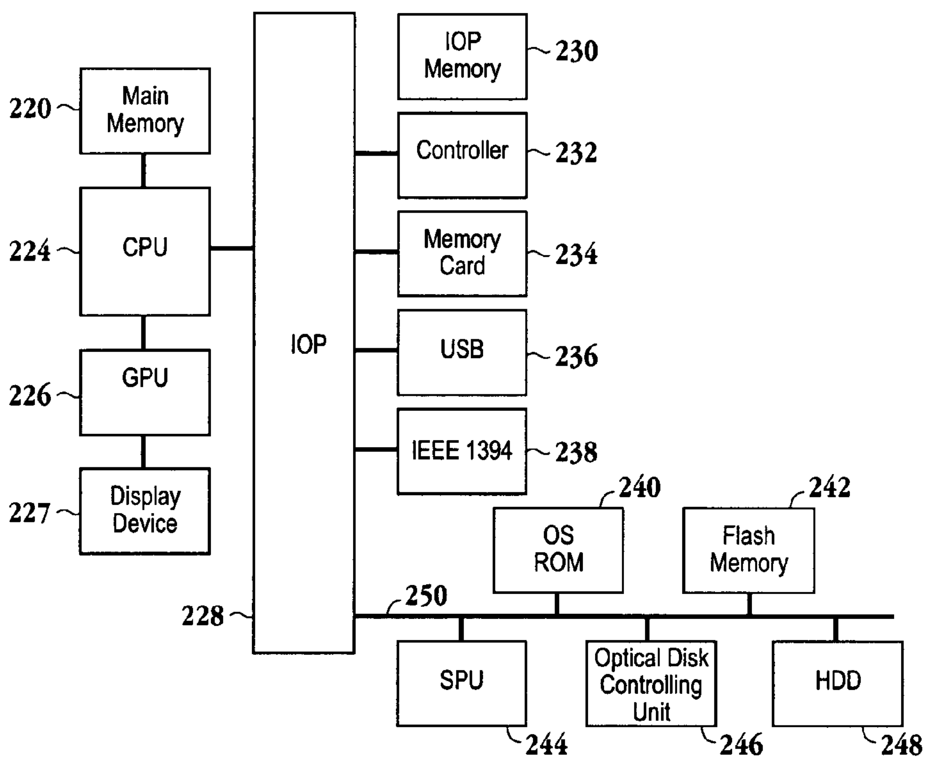


Fig. 10



220
Main Memory
224
CPU
226
GPU
227
Display Device
228
IOP
IOP Memory
230
Controller
232
Memory Card
234
USB
236
IEEE 1394
238
240
OS ROM
242
Flash Memory
250
SPU
244
Optical Disk Controlling Unit
246
HDD
248

Fig. 11

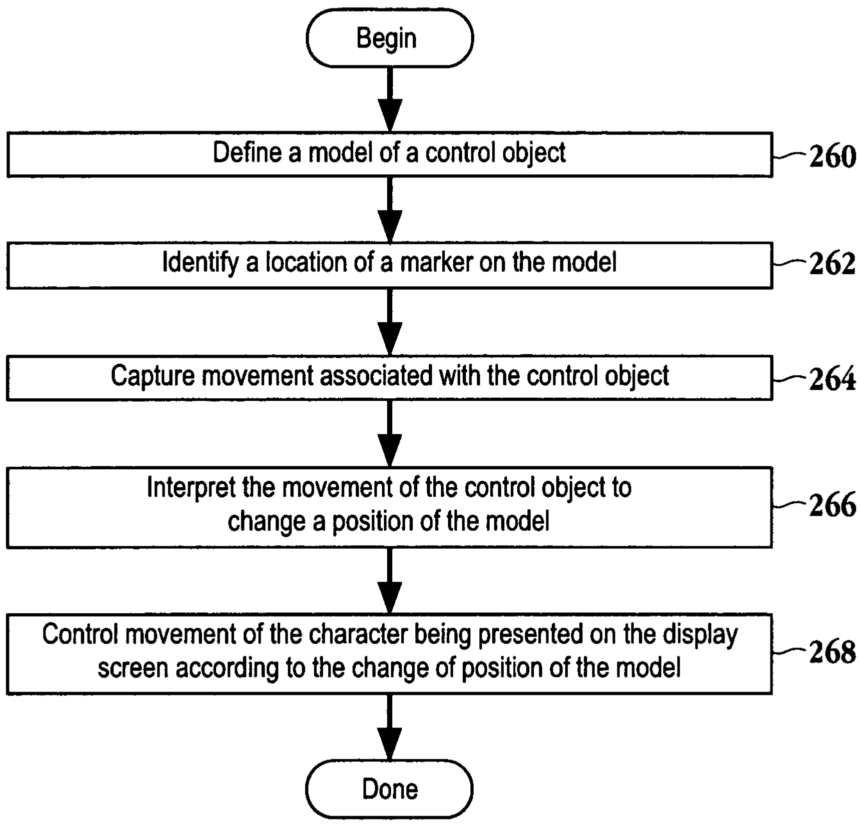

Begin
Define a model of a control object
260
Identify a location of a marker on the model
262
Capture movement associated with the control object
264
Interpret the movement of the control object to change a position of the model
266
Control movement of the character being presented on the display screen according to the change of position of the model
268
Done

Fig. 12

METHOD AND APPARATUS FOR REAL TIME MOTION CAPTURE

BACKGROUND OF THE INVENTION

1. Field of the Invention

This invention relates generally to video processing and more particularly to tracking depth of an image and/or a marker associated with the image to provide real time motion capture for video game applications.

2. Description of the Related Art

Movies, video games and other related video productions capture distance and dimension information. With respect to movies, an actor or sport figure's moves may be separately captured. For example, a sport figure may be filmed by a plurality of cameras while performing an activity. Objects, i.e., balls, attached to the sport figure enable the tracking of the moves. The tracked movements are then stored in a database for later use in a movie or video game. The movements may be played back by running through the frames of the stored movements in the database.

FIG. **1** is a schematic diagram of a person having a plurality of balls distributed over their body to store the persons moves. Person **100** has a plurality of balls **102** distributed over their body. Person **100** will then perform some activity which is captured by a plurality of cameras **104**. The captured video data may then be edited as desired to define video data **106**. Video data **106** is then stored in database **108** for later retrieval for insertion into a movie.

The scheme described with respect to FIG. **1** works well for a controlled environment, such as movie editing, where real time motion capture is not needed. In addition, since the person's movements are captured and stored, the person only has to wear the balls once. However, the motion capture as described with respect to FIG. **1** is used in a passive sense, i.e., editors insert the stored movement into a movie while editing the movie. Thus, the captured movement does not control any aspects of the movie and is inserted into a movie at a later time. The same holds true for a video game, i.e., captured motion is not used to control any aspects of a video game. In part, the complex computational capabilities and the need for point by point scanning has prohibited the use of captured motion to act as a control in real time.

As a result, there is a need to solve the problems of the prior art to provide a method and apparatus for providing real time motion capture that is capable of controlling aspects of a video game such as object and character movements.

SUMMARY OF THE INVENTION

Broadly speaking, the present invention fills these needs by providing a method and apparatus enabling control of a video game character and object through real time motion capture of a person playing the video game. It should be appreciated that the present invention can be implemented in numerous ways, including as a method, a system, computer readable media or a device. Several inventive embodiments of the present invention are described below.

In one embodiment, a method for real time motion capture for control of a video game character is provided. The method initiates with defining a model of a control object. Then, a location of a marker on the model is identified. Next, movement associated with the control object is captured. Then, the movement associated with the control object is interpreted to change a position of the model. Next, movement of the character being presented on the display screen is controlled according to the change of position of the model.

In another embodiment, a method for controlling an object presented on a display screen in communication with a computing device through real time motion capture is provided. The method initiates with identifying a depth image associated with an object being tracked. Then, a model associated with both an object presented on a display screen and the object being tracked is identified. Next, the model is fit to the depth image to capture motion associated with the object being tracked. Then, the object presented on the display screen is controlled in real time according to the fitting of the model to the depth image.

In yet another embodiment, a method for controlling movements of an image presented on a display screen through real time motion capture is provided. The method initiates with defining a model of a person. Then, a location of a marker on the model of the person is identified. Next, a depth image corresponding to a portion of the model of the person is provided. The portion of the model includes the marker. Then, the location of the marker is associated with a point on the depth image. Next, the portion of the model is positioned based upon a configuration of the depth image. Then, a video character displayed on a display screen is controlled according to the positioning of the portion of the model.

In still yet another embodiment, a computer readable media having program instructions for controlling an object presented on a display screen, in communication with a computing device, through real time motion capture is provided. The computer readable media includes program instructions for identifying a depth image of an object being tracked and program instructions for identifying a model associated with both an object presented on a display screen and the object being tracked. Program instructions for fitting the model to the depth image to capture motion associated with the object being tracked are included. Program instructions for controlling the object presented on the display screen in real time according to the fitting of the model to the depth image are provided.

In another embodiment, a system enabling control of an object through real time motion capture associated with a control image is provided. The system includes a computing device. A display screen in communication with the computing device is provided. The display screen is configured to display an image of an object, where data corresponding to the image of the object is provided to the display screen by the computing device. A video capture device in communication with the computing device is includes. The video capture device is enabled to track a control object. The video capture device is configured to translate motion associated with the control object to control motion of the object being displayed on the display screen as the control object moves.

In yet another embodiment, a system enabling control of video character through real time motion capture associated with a control image is provided. The system includes a computing device. Means for displaying an image of a video character from data received by the computing device and means for capturing a depth image associated with a control object are provided. Means for fitting a model to the depth image to define movement of the control object are included. Means for translating the movement of the control object to control motion associated with the video character on a display screen as the control object moves are also included.

Other aspects and advantages of the invention will become apparent from the following detailed description, taken in conjunction with the accompanying drawings, illustrating by way of example the principles of the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention will be readily understood by the following detailed description in conjunction with the accompanying drawings, and like reference numerals designate like structural elements.

FIG. **1** is a schematic diagram of a person having a plurality of balls distributed over their body to store the persons moves.

FIG. **2** is a simplified schematic diagram illustrating real time motion capture used to control a character of a video game in accordance with one embodiment of the invention.

FIG. **3** is a schematic diagram of a model of a person where the model includes markers depicting joint angles in accordance with one embodiment of the invention.

FIGS. **4**A through **4**C illustrate exemplary patterns that may be worn by the control object as a marker in accordance with one embodiment of the invention.

FIG. **5** is a schematic diagram of real time motion capture of a control object being used for controlling movement associated with a character presented on a display screen in accordance with one embodiment of the invention.

FIG. **6**A is a schematic diagram of a bodysuit having markers included in the bodysuit in accordance with one embodiment of the invention.

FIG. **6**B is an alternative embodiment to the bodysuit of FIG. **6**A.

FIG. **7** is a schematic diagram illustrating an optimization technique for monitoring an image frame for markers in accordance with one embodiment of the invention.

FIG. **8** is a schematic diagram of a model depicting regions filled in by a depth image in accordance with one embodiment of the invention.

FIG. **9** is a block diagram of an exemplary user input system for interaction with an object on a graphical display that can be used to implement embodiments of the present invention.

FIG. **10** illustrates an input system for user interaction with an object on a graphical display, according to embodiments of the present invention.

FIG. **11** is a simplified block diagram of a computer processing system configured to implement the embodiments of the invention described herein

FIG. **12** is a flowchart diagram of the method operations for real time motion capture for controlling a character presented on a display screen in accordance with one embodiment of the invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

An invention is described for a system, apparatus and method that enables real time motion capture which may be used to control a character or object of a video game. It will be obvious, however, to one skilled in the art, that the present invention may be practiced without some or all of these specific details. In other instances, well known process operations have not been described in detail in order not to unnecessarily obscure the present invention. FIG. **1** is described in the "Background of the Invention" section. The term about as used to herein refers to +/−10% of the referenced value.

The embodiments of the present invention provide a system and method for allowing real time motion capture to control a video game character or object. The movement of a person playing a video game, is used to control a character on a display screen associated with the video game. Thus, real time motion capture of the person playing the video game is used to control the character on the display screen. Here, the raw movement data of the person playing the video game is captured and used for control purposes in real time, as opposed to movement capture that is edited offline and then inserted into a movie at a later point in time.

In one embodiment, the movement of a human character presented on a display screen is controlled by the movement of a user or a person acting as a control object. The movement of the user may be captured through a video capture device, also referred to as an image capture device, such as a web cam or some other suitable type of camera. In one embodiment, the video capture device is configured to capture a depth image of the control object. The depth image of the control object provides data associated with a distance or depth associated with each pixel relative to the video capture device. Therefore, a skeleton associated with the control object and defining the location of joint angles may be mapped to a human character or even a non-human character. In another embodiment, markers worn by the control object may be used to indicate the location of joint angles. The depth images are combined with the skeleton using the locations defined by the markers to create a control object. In turn, motion of the control object is then used to control a character image on a display screen. For example, the character image may be an image from a video game. Thus, the person playing the video game is enabled to control the character image showing on a display screen. As will described in more detail below, the embodiments of the invention described herein, include real time motion capture for controlling a video character through markers. In another embodiment, a video capture device configured to provide depth information may be used with or without markers for the real time motion capture used to control a video character. As will described below, the markers may include colors, patterns, retro-reflective material, light, etc.

FIG. **2** is a simplified schematic diagram illustrating real time motion capture used to control a character of a video game in accordance with one embodiment of the invention. Motion capture input **120** is stored in motion capture database **124**. Control object motion **122** is captured through a capture device, such as a web cam, and associated with corresponding motion capture input **120** stored in database **124**. The associated motion capture input is then presented on a display screen as represented as output motion from data base **126**. For example, control object motion **122** may capture some motion or activity of a person playing a video game. In one embodiment, motion capture database **124** stores a plurality of motion capture input files **120** that capture the motion of a professional athlete. Accordingly, control object motion **122** is associated with a similar motion capture input **120** that corresponds to the control object motion and the similar motion capture input is presented on a display screen in real time. Therefore, the effect that the control object, e.g., a person playing a video game, is controlling the motion or activity of the professional athlete in real time is provided. It should be appreciated that motion capture input may also be artist animated data, e.g., non-human game characters, animated character images, etc. In another embodiment, control object motion **122** is used as constraint information. Here, the constraint information drives the animation of a character image, therefore, control object motion **122** directly controls the motion of a game character. It should be appreciated that a physical simulation system may accomplish this feature. In one embodiment, the game character mirrors the motion of the control object. In another embodiment, certain movement of the control object is mapped to cause a different movement of the game character. For example, when the control object moves a limb, arm or leg, the limb movement may be mapped to cause the eyebrows of a video game character to move. It should be appreciated that any type of control object movement may be mapped to any different type of movement of a character image, such as a video game character.

Still referring to FIG. **2**, in a more specific example, control object motion **122** may depict a person playing a basketball video game where the person performs a shooting motion as depicted in block **122***a*. The shooting motion of block **122***a* is associated with a shooting motion of a professional basketball player stored in database **124**. The shooting motion of the professional basketball player is displayed on a viewable screen as depicted in block **126***a*. Therefore, the motion capture of the person playing the video game is used to control, in real time, the professional character performing a similar motion. It should be appreciated that the basketball example is shown for illustrative purposes and not meant to be limiting. That is, the motion of the person playing the video game may correspond to any sport motion. Furthermore, the motion of the person playing the video game may be associated with non-sport activities. For example, the captured motion of the person playing the video game may be used to control movements of an animal or other non-human living object being displayed by the video game or even an inanimate object.

FIG. **3** is a schematic diagram of a model of a person where the model includes markers depicting joint angles in accordance with one embodiment of the invention. In one embodiment, the model is a skeleton image. As used herein, a skeleton image refers to any model of a structure that is tracked and is not limited to an articulated model where the model is rigid but has joints. Furthermore, the skeleton image may be defined with varying precision, e.g., a variety of joint constraints. Of course, the more joints and limbs associated with the skeleton image or the model, correlates to more data required to be tracked. Here, markers **140***a*-**140***i* are distributed over skeleton image **142**. Markers **140***a* and **140***d* correspond to the wrist location, markers **140***b* and **140***c* correspond to an elbow location while marker **140***e* corresponds to the torso. Markers **140***g* and **140***f* correspond to the knees and markers **140***h* and **140***i* correspond to the ankles. Of course, the embodiments described herein are not limited to the placement of the markers or the number of markers depicted in FIG. **3**, as more or less markers can be used. For example, headphones **141** can be used as a marker indicating a position of a head of the control object. One skilled in the art will appreciate that a headband may be used as a marker. Here, ear insertable devices may be provided where the ear insertable devices act as a marker and provide sound for the control object. Thus, a video capture device, e.g., camera, provides an image of a person playing the video game where the person playing the video game is wearing the markers. For example, the markers may be configured as straps a person can wear or the markers may be incorporated into the fabric of the person, i.e. control object. Software can then analyze the captured image of the person with the markers to create skeleton image **142** with known location of the joint angles as provided by markers **140***a*-**140***i*.

In one embodiment, the video capture device is configured to provide a depth image that can be used to fill in portion of the skeleton image and position the skeleton image in three dimensional space. Thus, markers **140***a*-**140***i* provide data as to a starting location of an appendage or limb and the depth image from the video capture device can fill in the appendage or limb in three dimensional space. As used herein, the terms appendage and limb are not meant to be limiting as a person, i.e., control object, may be controlling an object that is also captured by the depth image. Thus, the appendage or limb may include the objects being controlled by the person acting as a control object. In another embodiment, the video capture device does not have depth capturing capability and the markers will indicate a known location in space for a starting point of an appendage or limb. Here, an arm, hand, leg, foot, or some other appendage may be filled in from a database storing a typical configuration of the associated appendage. For example, a depth image of the control object can be taken prior to starting the game and the depth image may be stored in a database as a model to be used as needed for the control of a video game character in real time. In yet another embodiment, a video capture device is used to capture motion of the person playing the video game without markers. Here, certain assumptions about the location of the object, e.g., which hand, foot, etc., are from the right side or left side are made in order to translate the captured motion for control of a video game character. Thus, in each embodiment described above, the captured motion is used to control motion associated with a character of a video game in real time.

It should be appreciated that the markers may take on various forms. For example, material having a certain shape, color, pattern, reflective capability, or some other distinguishing quality so that a video capture device can identify a point in space by the marker may be used. One skilled in the art will appreciate that retro-reflective material may be used to provide a distinguishing reflective capability. Additionally, a light associated with the video capture device may be used in combination with the retro-reflective tape to provide the location in space of the starting point for a particular appendage of the control object. In one embodiment, the markers may take the form of blinking lights. Here, the lights could be non-visible light, such as infrared light. The lights may blink at a set frequency where the set frequency corresponds to a particular person or team.

FIGS. **4**A through **4**C illustrate exemplary patterns that may be worn by the control object as a marker in accordance with one embodiment of the invention. The exemplary patterns include a striped pattern of FIG. **4**A, a block pattern of FIG. **4**B and a crosshatching pattern of FIG. **4**C. It will be apparent to one skilled in the art that the patterns illustrated in FIGS. **4**A-**4**C are not meant to be limiting as any type of pattern capable of being recognized by a video capture device may be used. It should be appreciated that the different patterns may be used to differentiate between right and left limbs in one embodiment. For example, marker **140***d* of FIG. **3** may include the pattern of FIG. **4**A, while marker **140***a* of FIG. **3** may include the pattern of FIG. **4**B. Accordingly, the right and left wrists will be differentiated. Additionally, shapes, colors, etc., may be used in conjunction with the patterns to define a marker. Thus, should the pattern become deformed through movement, a backup indicator is still available for the marker.

FIG. **5** is a schematic diagram of real time motion capture of a control object being used for controlling movement associated with a character presented on a display screen in accordance with one embodiment of the invention. Skeleton image **142**, corresponding to a person playing a video game, is tracked by video capture device **148**. In one embodiment, video capture device **148** is configured to capture depth data for each pixel. For example, depth cameras offered by 3DV STSTEMS or CANESTA may be employed as video capture device **148** for the embodiments described herein that capture depth data. Where the person playing the video game is wearing retro-reflective tape as a marker, camera **148** includes light **150**. In one embodiment, light **150** may be an infrared light. In another embodiment, light **150** is blinking so that it is possible to differentiate between permanent lights in a scene and markers to be tracked, i.e., the reflection of the blinking light from the retro-reflective tape. The retro-reflective tape, reflects all of the light from light **150**, therefore, the retro-reflective tape will be a bright spot in the image data provided by video capture device **148**. The bright spot or void will indicate a starting point for an appendage associated with the marker. For example, marker **140***d* is associated with a right wrist and would indicate the starting point for a depth image including the right wrist, which could include a hand or a forearm. Of course, the hand could be controlling an object as described above. It should be appreciated that camera **148** may include several blinking lights of different colors. Here, the retro-reflective tape, or any suitable reflective marker, may be designed to reflect different colors, thereby allowing for differentiation between a number of reflective markers.

Marker **140***c* of FIG. **5** indicates the position of the torso. As mentioned above, any number of markers may be worn by the control object at defined locations. It will be apparent to one skilled in the art that the capture device can identify the particular marker as being associated with a particular appendage, i.e., wrist or torso with respect to FIG. **5**. Where video capture device **148** is configured to provide a depth image, the video capture device provides the data to fill in the rest of the body for skeleton image **140** with the depth information as described with reference to FIG. **8**. Camera **148** includes microprocessor **152** configured to combine the depth information with the marker location to generate and track the image of the control object, i.e., person playing the video game. Microprocessor **152** may perform processing functions as described with reference to FIG. **9**. Camera **148** is in communication with computing device **154**, which in turn is in communication with display device **156**. In one embodiment, computing device **154** is a video game console such as the "PLAYSTATION 2"® manufactured by Sony Computer Entertainment Inc. In another embodiment, light **150** may be configured to emit light in the infrared spectrum to minimize the visible light directed at the control object.

FIG. **6**A is a schematic diagram of a bodysuit having markers included in the bodysuit in accordance with one embodiment of the invention. Markers **162***a* through **162***i* may be integrated into bodysuit **160** at joint locations corresponding to the locations discussed with reference to FIG. **3**. In one embodiment, torso marker **162***e* may be encoded with data that identifies the person wearing the suit with a certain sports figure, celebrity, team identification, etc. Additionally, a marker may be encoded to provide the person wearing the suit with enhanced features during a video game. For example, the bodysuit may provide the user extra protection through more body armor or other enhancements that would entice consumers to purchase the bodysuits for use with a video game.

FIG. **6**B is an alternative embodiment to the bodysuit of FIG. **6**A. Here, micro fibers of retro-reflective material are woven into bodysuit, i.e., incorporated throughout the fabric of bodysuit **160**. Thus, the retro-reflective material is distributed throughout the bodysuit. Here, the body suit would appear as a maximum to a depth camera, but would be preferable for the embodiment where the camera is not enabled to capture depth data. That is, where the video capture device is configured to provide two dimensional image data, it is preferable to have as much of the control object to be as reflective as possible, in order to more easily track the control object for control of a video game character in real time. Of course, it is not necessary for the control object to be covered by reflective material as strategically placed straps will suffice as discussed above.

FIG. **7** is a schematic diagram illustrating an optimization technique for monitoring an image frame for markers in accordance with one embodiment of the invention. Here, region **170** which defines an area around skeleton image **142**. Thus, rather than search the entire frame of data for a marker, only a portion of the frame defined around skeleton image **142** is searched for successive frames. Here, a portion of frames **174-2** through **174-***n* would be searched to enhance performance. Skeleton image **142** is shown controlling object **172**, where object **172** is a sword. It should be appreciated that while the control object has mainly been referred to as a person, the person may control an object such as a sword, ball, bat, etc. Thus, the motion associated with the object being controlled may also be captured to control a similar object on a display screen in real time, e.g., where the video character is controlling a sword.

FIG. **8** is a schematic diagram of a model depicting regions filled in by a depth image in accordance with one embodiment of the invention. As mentioned with reference to FIG. **3**, a video capture device, e.g., camera, provides an image of a person playing the video game where the person playing the video game is wearing the markers. The captured image of the person is analyzed with the markers to create skeleton image **142**, also referred to as a model, with known location of the joint angles as provided by markers **140***a*-**140***i*. In one embodiment, the person wearing the markers stands in front of the video capture device to define a model or body mesh of the control object that is stored as mentioned above. That is, a calibration step is performed initially to define a model depth image for the control object. Then, the depth image or portions of the depth image are used to fill in skeleton **142** to create a three dimensional image. The respective marker defines a position for the corresponding depth image portion to be located. For example, left and right forearms are be defined in regions **180***c* and **180***e*, respectively. Upper torso and lower torso are defined in regions **180***d* and **180***b*, respectively, while left shin and right shin are defined in regions **180***a* and **180***f*, respectively. Additionally, hands and feet are defined below respective markers **140***a*, **140***d*, **140***i* and **140***h*. Thus, the markers define a starting location and the depth image is used to fill in the remainder to provide a three dimensional control object for controlling the motion of a video game character in real time.

FIG. **9** is a block diagram of an exemplary user input system for interaction with an object on a graphical display that can be used to implement embodiments of the present invention. As shown in FIG. **9**, the user input system is comprised of a video capture device **200**, an input image processor **202**, an output image processor **204**, and a video display device **206**. Video capture device **200** may be any device capable of capturing sequences of video images, and, in one embodiment, is a digital video camera (such as a "web-cam"), or similar image capturing device. As mentioned above, the video capture device may be configured to provide depth image. Input image processor **202** translates the captured video images of the control object into signals that are delivered to an output image processor. In one embodiment, input image processor **202** is programmed to isolate the control object from the background in the captured video image through the depth information and generate an output signal responsive to the position and/or movement of the control object. The output image processor **106** is programmed to effect translational and/or rotational movement of an object on the video display device **108** in response to signals received from the input image processor **104**.

These and additional aspects of the present invention may be implemented by one or more processors which execute software instructions. According to one embodiment of the present invention, a single processor executes both input image processing and output image processing as illustrated in FIG. **5**. However, as shown in the figures and for ease of description, the processing operations are shown as being divided between an input image processor **202** and an output image processor **204**. It should be noted that the invention is in no way to be interpreted as limited to any special processor configuration, such as more than one processor. The multiple processing blocks shown in FIG. **9** are shown only for convenience of description.

FIG. **10** illustrates an input system for user interaction with an object on a graphical display, according to embodiments of the present invention. The input system environment includes control object **212**, video capture device **214**, video display device **216**, and console **208** containing the processor functionality, such as a video game machine. Control object **212** in the input system environment should be located within the field of view **210** of the video capture device **214**. Processing system **208** can be implemented by an entertainment system or game console, such as a Sony® Playstation™ II or Sony® Playstation™ I type of processing and computer entertainment system. It should be noted, however, that processing system **208** can be implemented in other types of computer systems, such as personal computers, workstations, laptop computers, wireless computing devices, or any other type of computing device that is capable of receiving and processing graphical image data. Of course, control object **212** may incorporate markers as described above and/or video capture device **214** may include depth capturing capability.

FIG. **11** is a simplified block diagram of a computer processing system configured to implement the embodiments of the invention described herein. The processing system may represent a computer-based entertainment system embodiment that includes central processing unit ("CPU") **224** coupled to main memory **220** and graphical processing unit ("GPU") **226**. CPU **224** is also coupled to Input/Output Processor ("IOP") Bus **228**. In one embodiment, GPU **226** includes an internal buffer for fast processing of pixel based graphical data. Additionally, GPU **226** can include an output processing portion or functionality to convert the image data processed into standard television signals, for example NTSC or PAL, for transmission to display device **227** connected external to the entertainment system or elements thereof. Alternatively, data output signals can be provided to a display device other than a television monitor, such as a computer monitor, LCD (Liquid Crystal Display) device, or other type of display device.

IOP bus **228** couples CPU **224** to various input/output devices and other busses or device. IOP bus **228** is connected to input/output processor memory **230**, controller **232**, memory card **234**, Universal Serial Bus (USB) port **236**, IEEE1394 (also known as a Firewire interface) port **238**, and bus **250**. Bus **250** couples several other system components to CPU **224**, including operating system ("OS") ROM **240**, flash memory **242**, sound processing unit ("SPU") **244**, optical disc controlling unit **246**, and hard disk drive ("HDD") **248**. In one aspect of this embodiment, the video capture device can be directly connected to IOP bus **228** for transmission therethrough to CPU **224**; where, data from the video capture device can be used to change or update the values used to generate the graphics images in GPU **226**. Moreover, embodiments of the present invention can use a variety of image processing configurations and techniques, such as those described in U.S. patent application Ser. No. 09/573,105 filed May 17, 2000, and entitled OUTLINE GENERATING DATA, GENERATING METHOD AND APPARATUS, which is hereby incorporated by reference in its entirety.

Programs or computer instructions embodying aspects of the present invention can be provided by several different methods. For example, the user input method for interaction with graphical images can be provided in the form of a program stored in HDD **248**, flash memory **242**, OS ROM **240**, or on memory card **232**. Alternatively, the program can be downloaded to the processing unit through one or more input ports coupled to CPU **224**. The program modules defining the input method can be provided with the game or application program that is executed by CPU **224** and displayed on display device **227** or they may be provided separately from the application program, such as for execution from local main memory **220**.

Embodiments of the present invention also contemplate distributed image processing configurations. For example, the invention is not limited to the captured image and display image processing taking place in one or even two locations, such as in the CPU or in the CPU and one other element. For example, the input image processing can just as readily take place in an associated CPU, processor or device that can perform processing; essentially all of image processing can be distributed throughout the interconnected system. Thus, the present invention is not limited to any specific image processing hardware circuitry and/or software. The embodiments described herein are also not limited to any specific combination of general hardware circuitry and/or software, nor to any particular source for the instructions executed by processing components.

FIG. **12** is a flowchart diagram of the method operations for real time motion capture for controlling a character presented on a display screen in accordance with one embodiment of the invention. The method initiates with operation **260** where a skeleton image or a model of a control object is defined. The control object may be a person playing a video game associated with the character presented on the display screen. It should be appreciated that the person may stand in front of a video capture device within the field of view of the video capture device in order to define a model of the person as described above. Thus, the skeleton image or model may be defined through a self calibration process where the person stands in front of the video capture device. The method then advances to operation **262** where a location of a marker on the skeleton image or model is identified. For example, the markers may indicate the location of joint angles on the skeleton image or model as described with reference to FIGS. **3**, **5** and **8**. It should be appreciated that the markers may take on the various formats as described herein.

The method of FIG. **12** then proceeds to operation **264** where movement associated with the control object is captured. In one embodiment, the movement is captured through a camera configured to provide a depth image, i.e., image data associated with a z axis as well as an x axis and y axis. In another embodiment, the movement is captured in two dimensions by a digital video camera, such as a web cam. The method then moves to operation **266**, where in response to the movement of the control object, a position of the skeleton image or model is changed. That is, the skeleton image is moved to correspond with the movement of the control object. It should be appreciated that where a depth camera is used as a video capture device, a depth image corresponding to the skeleton image may be provided here. The skeleton image is positioned to correspond with the position of the depth image.

For exemplary and illustrative purposes, the skeleton image may be thought of as a rag doll that is positioned to correspond with the position of the depth image. As the control object moves, the depth image tracks the movement of the control object, in turn, the skeleton image is also similarly moved to follow the depth image. Thus, the movement of the control object is repeated by the skeleton image. As mentioned above, the markers may be used without the depth camera. Similarly, the depth camera may be used without the markers. It should be appreciated that the depth camera captures the three dimensional data, however, the depth camera does not have the knowledge to map the three dimensional data to the skeleton. The markers provide known data points in space, therefore, the markers enable the depth data to be mapped to the skeleton. Of course, without the use of the markers certain assumptions may be made so that the depth data can be mapped to the skeleton. That is, the markers eliminate the need for the assumptions as an actual point, such as a joint angle, is labeled by the marker and the corresponding depth data can be filled in from the point identified by the marker. The method then advances to operation **268** where the movement of the character being presented on the display screen is controlled according to the change of position of the skeleton. Here, the movement of the character is being controlled in real time. That is, as the control object moves, the skeleton is positioned accordingly through the use of image data having depth information and/or markers.

In one embodiment, a system enabling control of an object through real time motion capture associated with a control image consisting of various components defined in terms of the functions of each of the components. The system includes a computing device, e.g., a game console. The system may include means for displaying an image of a display object from data received by the computing device. Here, the means for displaying an image mat be any suitable display screen including a television monitor. Means for capturing a depth image associated with a control object are included. In one embodiment, the means for capturing a depth image may be provided by a camera enabled to provide depth data, such as the cameras available from 3DV SYSTEMS or CANESTA mentioned above. Means for fitting a skeleton image or model to the depth image to define movement of the control object are provided. For example, a microprocessor may provide the means for fitting the skeleton image to the depth image to define movement of the control object. As mentioned above, the processing may be provided through one or more microprocessors. Means for translating the movement of the control object to control motion associated with the display object on a display screen as the control object moves. Here again a microprocessor, such as a graphics processing unit, can accomplish the means for translating the movement of the control object to control motion associated with the display object. Of course, the microprocessors for performing the above described functionality can be included in a chipset.

In summary, the above described invention describes a method and a system for providing real time motion capture for controlling a character of a video game. The embodiments described above allow for real time motion capture which may be used to control a character or object of a video game. The ability to track depth combined with a marker can provide a starting point for identifying the limbs/body of a person (control object) being tracked. The person can then be filled in to resemble a true character or person. Thus, once a depth image is identified, a skeleton image associated with the depth image can be filled in from a database having the depth images or from depth images captured through the video capture device. Various forms of markers may be used in conjunction with a video capture device that may or may not capture depth images. Examples can include localized placement of wearable retro-reflective tape, e.g., on the wrists, knees, head, etc. Alternatively, body suits or clothes incorporating different types of markers (patches, identifiable thread, etc) may be provided. In one embodiment, once movement is identified (or detected), a reference to a database of movements may locate a corresponding movement in the database. For example, if the user of game selects a particular character, such as a famous ball player, the movement may be like that performed by the actual famous ball player. These movements would be stored in the database. The obtained database movement will thus be used to control the character (person/skeleton) that is part of the game. In one embodiment, a camera or hardware configured to track the depth can include a processor, special circuitry or a DSP that is configured to do the filling in, the database look ups and translate the user's actions into "on-screen" activity. As discussed above the term skeleton image may be defined broadly to include any model of any object being tracked, whether the object is a human object, an animated object or an inanimate object.

With the above embodiments in mind, it should be understood that the invention may employ various computer-implemented operations involving data stored in computer systems. These operations include operations requiring physical manipulation of physical quantities. Usually, though not necessarily, these quantities take the form of electrical or magnetic signals capable of being stored, transferred, combined, compared, and otherwise manipulated. Further, the manipulations performed are often referred to in terms, such as producing, identifying, determining, or comparing.

The above described invention may be practiced with other computer system configurations including hand-held devices, microprocessor systems, microprocessor-based or programmable consumer electronics, minicomputers, mainframe computers and the like. The invention may also be practiced in distributing computing environments where tasks are performed by remote processing devices that are linked through a communications network.

The invention can also be embodied as computer readable code on a computer readable medium. The computer readable medium is any data storage device that can store data which can be thereafter read by a computer system. Examples of the computer readable medium include hard drives, network attached storage (NAS), read-only memory, random-access memory, CD-ROMs, CD-Rs, CD-RWs, magnetic tapes, and other optical and non-optical data storage devices. The computer readable medium can also be distributed over a network coupled computer system so that the computer readable code is stored and executed in a distributed fashion.

Although the foregoing invention has been described in some detail for purposes of clarity of understanding, it will be apparent that certain changes and modifications may be practiced within the scope of the appended claims. Accordingly, the present embodiments are to be considered as illustrative and not restrictive, and the invention is not to be limited to the details given herein, but may be modified within the scope and equivalents of the appended claims. In the claims, elements and/or steps do not imply any particular order of operation, unless explicitly stated in the claims.

What is claimed is:

1. A method for real time motion capture for control of a video game character during game play, comprising:

defining a model of a control object, wherein the model is mapped to the video game character;

identifying a marker on the control object and mapping the marker to the model;

associating a location of the marker to the model;

capturing movement associated with the control object with a single video capture device;

interpreting the movement associated with the control object to change a position of the model based on data captured through the single video capture device, the data captured including the marker, in the interpreting, for certain movements during game play, associating in real-time, the movement associated with the control object to a corresponding motion capture input stored in a database, the corresponding motion capture input is predefined to be used without modification and performed by the video game character, the predefined motion of the video game character is presented in real time and in response to a similar movement of the control object;

for certain movements during game play, the movements associated with the control object are mirrored to the video game character without use of the predefined motion; and

movement of the video game character being presented on the display screen in real time according to the change of position of the model.

2. The method of claim **1**, wherein the control object is a person and the model is a skeleton image of the person.

3. The method of claim **1**, wherein the method operation of capturing movement associated with the control object includes,

capturing movement associated with an object being controlled by the control object.

4. The method of claim **1**, wherein the method operation of interpreting the movement associated with the control object to change a position of the model includes,

accessing a database having depth images corresponding to the model.

5. The method of claim **1**, further comprising:

continuing to capture movement associated with the control object, interpret the movement associated with the control object to change a position of the model and control movement of the character being presented on the display screen according to the change of position of the model.

6. A method for controlling an object presented on a display screen in communication with a computing device through real time motion capture during video game play, comprising:

capturing movement of a depth image associated with an object being tracked through a single video capture device;

identifying a model associated with both an object presented on a display screen and the object being tracked,

after one or more frames are captured by the single capture device, performing operations of,

fitting the model to the depth image to captured motion associated with the object being tracked through data from the single video capture device, the fitting includes filling a three dimensional space of the model for the object presented on the display screen that corresponds to the object being tracked;

during the video game play,

for certain movements being tracked, associating the movement of the object being tracked to a corresponding stored motion capture input stored in a database, wherein the corresponding motion is a predefined motion to be used without modification and performed by the object presented on the display screen, the predefined motion is in response to a similar movement of the object being tracked;

for certain movements being tracked, mirroring the movement of the object being tracked without use of the predefined motion;

controlling the object presented on the display screen in real time according to the movement of the object being tracked; and

displaying the corresponding stored motion capture input on the display screen in real time.

7. The method of claim **6**, wherein the object being tracked is a person interacting with the computing device and the model is a skeleton image of the person.

8. The method of claim **6**, wherein the computing device is a video game console.

9. The method of claim **6**, wherein the object is presented on a display screen is a video game character.

10. The method of claim **6**, wherein the object being tracked does not include markers.

11. The method of claim **1**, wherein associating real time movement of the control object to a corresponding motion capture input includes associating a captured control object motion with the corresponding predefined motion of the video game character stored in the database that mirrors the control object motion.

12. The method of claim **11**, wherein the control object motion drives the presentation of the corresponding predefined motion of the video game character that mirrors the control object motion in real time.

13. The method of claim **6**, wherein associating the movement of the object being tracked to a corresponding motion capture input includes associating a captured control object motion with the corresponding predefined motion that mirrors the movement of the object being tracked to be performed by the object presented on the display screen.

14. The method of claim **13**, wherein the control object motion drives the presentation of the corresponding predefined motion of the object presented on the display screen that mirrors the movement of the object being tracked.

* * * * *